Computer Architecture

Spring 2016

Final Review

Shuai Wang

Department of Computer Science and Technology

Nanjing University

Computer Architecture

- Computer architecture, like other architecture, is the art of determining the needs of the user of a structure and then designing to meet those needs as effectively as possible within economic and technological constraints.

---Frederick P.Brooks Jr, *Planning a Computer System: Project Stretch*, 1962

Performance

$$\text{CPU time} = \frac{\text{Seconds}}{\text{Program}} = \frac{\text{Instructions}}{\text{Program}} \times \frac{\text{Cycles}}{\text{Instruction}} \times \frac{\text{Seconds}}{\text{Cycle}}$$

AMAT (Average Memory Access Time) = Hit Time + Miss Rate x Miss Penalty

Amdahl's Law:

$$\text{Speedup} = 1 / ((1 - \text{Fraction}_{\text{enhanced}}) + \text{Fraction}_{\text{enhanced}} / \text{Speedup}_{\text{enhanced}})$$

Instruction Level Parallelism (ILP)

- Pipelining
- Speculation & Branch Prediction
- Out-of-Order Execution
- Dynamical Scheduling
- ……..

Pipelining

- Basic processor pipelining stages

- Pipelining hazards
  - Structural
  - Data
  - Control

- Data dependencies
  - RAW
  - WAW
  - WAR

Branch Prediction

- Dynamic vs. Static
- Direction vs. Target prediction
- BTB vs. RAS
- Branch History Table (BHT) vs. Pattern History Table (PHT)
  - Global history vs. Local history

Virtual Memory & Caches

- Page Table, TLB, & Cache
  - Address translate + Data access
  - Virtually indexed, physically tagged caches

- Cache
  - Principle of Locality
  - Direct-mapped, Fully Associative, Set Associative
  - Cache Block (Cacheline), Set, Tag, Way
  - Replacement Policy
    - FIFO, LRU, Random
  - Write through vs. Write Back
  - Cache Miss (3 C)
  - Optimization methods

OoO Execution & Dynamic Scheduling

- Scoreboarding (CDC 6600)
- Tomasulo's Algorithm (IBM 360/91)
- Tomasulo With Reorder Buffer
  - P6 style vs. MIPS R10K

Scoreboarding

- Implementation
  - Instructions dispatched in-order to functional units provided no structural hazard or WAW
  - Instructions wait for input operands (RAW hazards) before execution
  - Instructions wait for output register to be read by preceding instructions (WAR)
- Stages
  - Issue, Read operands, Execution, & Write Result
- In-order issue; out-of-order execute & commit

Tomasulo's Algorithm

- Similar to scoreboarding, but added renaming in hardware
  - Registers renamed to Reservation Stations (RS)
  - Common data bus (CDB)
  - Eliminate WAR/WAW hazards

- Stages
  - Issue, Execute, & Write Result

- In-order issue, out-of-order execution & commit.

Tomasulo + ROB

- Add reorder buffer (ROB) to fix speculation & precise exception
  - in-order commit
- Stages
  - Issue, Execution, Write result, & Commit
- In-order issue, out-of-order execution & writeback, in-order commit

MIPS R10K & Register Renaming

- P6 style vs. MIPS R10K
  - separate control (ROB/RS) from data (registers/FU)

- MIPS R10K
  - one big physical register file (PRF) holds all data → no copies
  - ROB and RS used only for control and tags → small

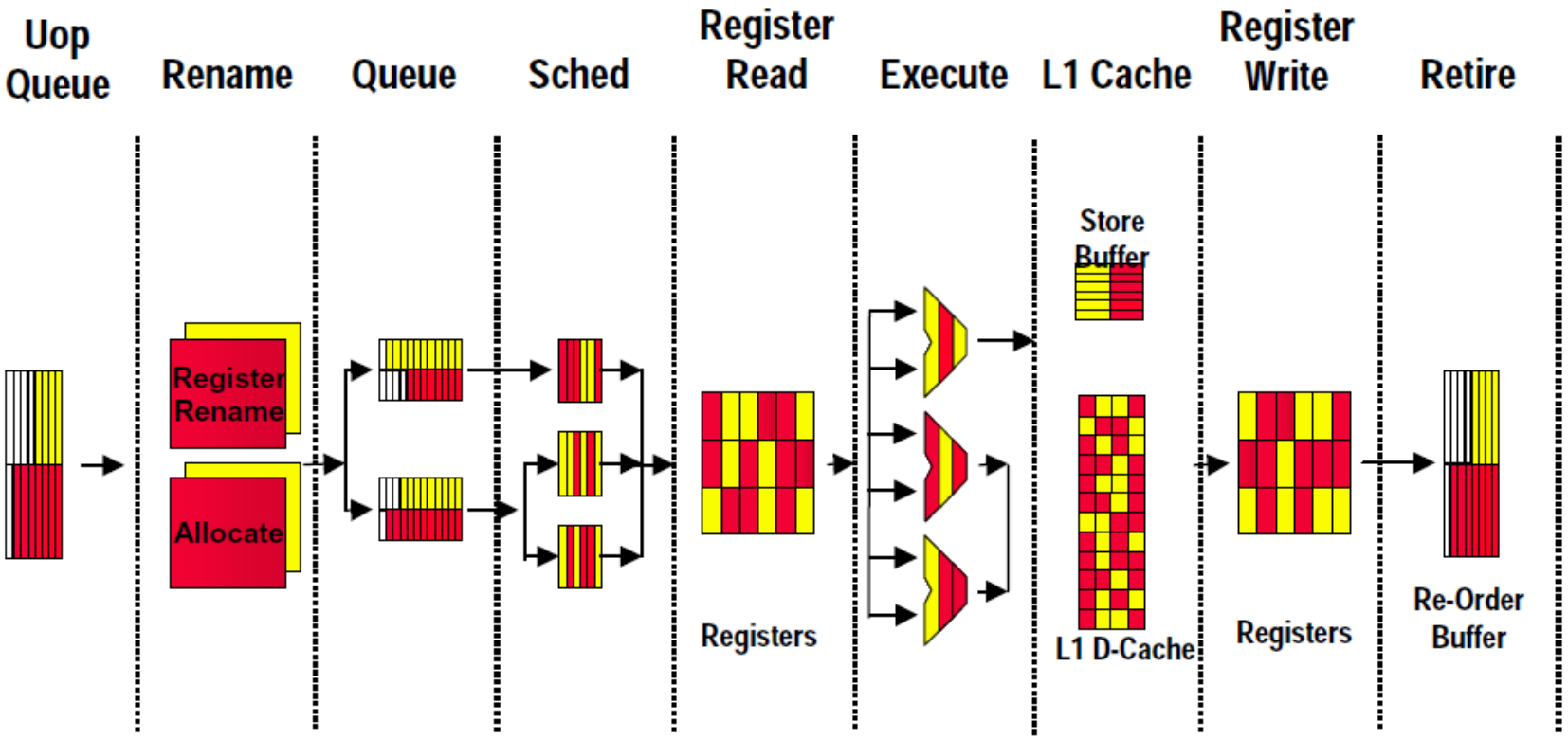
- Register Renaming
  - architectural register -> physical register
  - When to free physical register?

Simultaneous Multithreading (SMT)

- Dynamically scheduled processor already has many HW mechanisms to support multithreading
  - large set of registers that can be used to hold the register sets of independent threads (assuming separate renaming tables are kept for each thread)
  - out-of-order completion allows the threads to execute out of order, and get better utilization of the HW

- Just adding a per thread renaming table and keeping separate PCs
  - Independent commitment can be supported by logically keeping a separate reorder buffer for each thread

Pentium-4 Hyperthreading Execution Pipeline



Static Scheduling & VLIW

- Loop Unrolling

- Software Pipelining

Memory

- DRAM access
  - Activate
  - Read/write
  - Precharge

- Organization
  - Channel
  - DIMM
  - Rank
  - Chip
  - Bank
  - Row/Column

Multiprocessors

- SMP vs. CMP

- On-chip Interconnects
  - Bus
  - Crossbar
  - Ring
  - Mesh
  - Torus

Cache Coherence & Memory Consistency

- Snooping based vs. Directory based

- MSI Protocol
  - Invalid
  - Shared
  - Modified

- Sequential Consistency (SC)
  - Relaxed Consistency

Power & Reliability

- Dynamic and Static Power
  - Clock gating
  - Power gating
  - DVFS
- Hardware Faults & Soft Errors
  - Metrics
    - MTTF, MTTR, FIT….
  - Soft errors & AVF